

UNITED STATES PATENT APPLICATION
FOR

**METHOD AND APPARATUS FOR PROCESSING INTERNET
SITE NAMES THROUGH REGULAR EXPRESSION COMPARISON**

Inventor:

DEAN HILLER

Prepared By:

Kenyon & Kenyon
333 West San Carlos Street
San Jose, CA 95110
(408) 975-7500

Express Mail No.
EL372084900US

**Method and Apparatus for Processing Internet Site Names
through Regular Expression Comparison**

Field of the Invention

5 The present invention pertains to network communications. More particularly, the present invention pertains to the processing of network addresses and site names (such as those found in the Internet) using regular expressions.

Background of the Invention

10 A network system includes a number of computer systems and local networks connected in such a manner to allow communication between them. The Internet, which includes the so-called World Wide Web, is an example of a global network system. A computer system coupled to the Internet (e.g., via a dial-up connection to an Internet Service Provider (ISP)) may communicate with any of the computer systems coupled in the Internet system. In other words,

data, such as a request for information, may be directed to a particular computer system on the Internet.

Because of the number of computer systems in the network, a system is present to properly route data communicated by one computer system to another. This is assisted in the Internet system with a plurality of Domain Name Servers. These servers communicate directly with a subset of the servers that are coupled in the Internet system. When a first computer seeks to send data intended for a second computer, a request is made to a first Domain Name Server to find the address of the second computer system on the Internet. As is known in the Internet art, the second computer system is identified by a unique site name such as something having the form "abrainc.com." The first Domain Name Server that receives the request from the first computer system checks its stored site name information to see if there is an exact match for the "abrainc.com" character stream. If there is not a match, the first DNS sends a responsive message back to the first computer system. If there is a match, the first DNS forwards the request to another DNS where a similar comparison between the "abrainc.com" string and those strings stored at the DNS. If there is a match at this DNS, associated with the match is a detailed Internet address (typically a string of numbers and periods) that is sent back to the first computer system. After the Internet address is supplied to the first computer system, the first computer system then sends its data, request, etc. to the second computer. Through this system, data, requests, etc. are forwarded between nodes in the Internet until it reaches its destination.

In actuality, each unique Internet site name has associated with it a numerical Internet address. The DNS system is a hierarchical system. There are a small number of root DNSs which are responsible for tracking the top level domains such as .com, .net, .edu, etc. These

servers are instrumental in identifying valid site names. The system for locating properly addressed computer systems is a very robust system. The goal of the system is to forward requests to server systems that are more knowledgeable about the appropriate location of the DNS. The information that is stored in the DNSs is repeatedly updated so that future requests to a particular site name can be properly handled.

One problem associated with this system is in how it handles similar site names (i.e., site names that have a similar number of characters or have character strings that are in common). In the art, similar site names are handled in exactly the same manner as for every other site name. In other words, though similar site names may be directed to the same computer system, the DNS system must be updated for each and every site name to the computer system. This leads to inefficiency, in that it takes some time for the DNS system to be updated with the correct information so that newly created site names can be quickly located for communication.

Summary of the Invention

This and other problems are addressed by the method and apparatus of the present invention. According to an embodiment of the present invention a Domain Name Server is modified so as to use a regular expression in the comparison between the requested site name and the series of similar site names (whether actual or potential) associated with a computer system.

Brief Description of the Drawings

Fig. 1 is a block diagram of a computer network system operated according to an embodiment of the present invention.

Fig. 2 is a flow diagram of a method for operating a computer system of Fig. 1 according to an embodiment of the present invention.

Detailed Description

Referring to Fig. 1, a computer network that may be operated according to an embodiment of the present invention is shown. In this system, a first computer system 11 seeks to communicate with a second computer system 19. In doing so, it seeks to send some sort of information which will be referred to herein as “data” to a first site name which may be associated with computer system 19. In this example, the first site name is similar to a plurality of other site names and is prone for searching using a regular expression for comparison purposes as described in further detail below. For the purposes of this embodiment, the site name is as follows: 4085551234.abrainc.com (e.g., a U.S. telephone number with an “abrainc” sub-domain and a “.com” domain).

The first site name is communicated, for example, by the first computer system 11 to an Internet Service Provider (ISP) 13. The ISP 13 then transfers the site name to a Network Server 15 (e.g., over Internet 14). As known in the art, this Network Server may be a root Domain Name Server which is responsible for knowing valid site names in the “.com” realm. Once Network Server 15 identifies a valid “.com” site name (i.e., the “abrainc.com” portion of the site name), the site name is eventually forwarded to the Internet and preferably is forwarded to

Domain Name Server 17 to determine if 4085551234.abrainc.com is a valid site name (e.g., via a network node 18).

Referring to Fig. 2, a method of the present invention is shown in flow diagram form. In block 21, a normal name lookup is performed where the first site name is compared to the site names that are stored at the DNS 17. At decision block 23 it is determined whether a match has been found. If one has, control passes to block 24 where an appropriate response is sent by the DNS 17 to first computer system 11. If no match is found, control passes to decision block 25 where it is determined whether regular expression comparisons are enabled at the DNS 17. If they are not, then control passes to block 26 where an "unfound" message or the like is sent back to first computer system 11 (Fig. 1).

If regular expression comparisons are enabled, control passes to block 27 where a first regular expression is retrieved (e.g., from memory at DNS 17). According to an embodiment of the present invention, the first regular expression can have a form such as $\wedge d\{10\}$.X.Y where $\wedge d\{10\}$$ represents a string of ten numbers, X represents a sub-level domain ("abrainc" in this example), and Y represents a top-level domain ("com" in this example). An alternative regular expression could be $\wedge [0-9]+$.X.Y$ where $\wedge [0-9]+$$ represents a string of numbers (each between 0 and 9). In block 29 a comparison is made between the first site name and the regular expression. Such a comparison can be done, for example using a comparison procedure such as that which is used in the UNIX operating system. In this embodiment, the regular expressions set forth above are in a UNIX format. One skilled in the art based on the above teaching will appreciate that other formats may be used to represent a regular expression. In decision block 31, it is determined whether there is a match. If there is, then control passes to block 33 where a$

good response is sent to first computer system 11 with an appropriate Internet Protocol (IP) address. If there is not a match, then control passes to decision block 35 to determine whether other regular expressions exist. If there are no additional regular expressions present at DNS 17, then control passes to block 26 where an "unfound" message or the like is sent to computer system 11. If there are additional regular expressions, then control passes to block 37 to retrieve the next regular expression and pass control to block 29 to perform the next comparison.

Although several embodiments are specifically illustrated and described herein, it will be appreciated that modifications and variations of the present invention are covered by the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention. For example, Network Server 15 as a Domain Name Server may be modified to include a regular expression comparison as described above with reference to Fig. 1. In addition, other Domain Name Servers may be modified to include the regular expression method described above.

For example, the root Domain Name Server for a particular country may be used to perform the regular expression comparison. Accordingly, assuming that Z is a valid country code (i.e., a geographically-oriented top-level domain), all site names including that root are directed to the Domain Name Server for that country. Thus, modifying the regular expression above to $\wedge d\{10\}$.Z would result in the data being sent to the root Domain Name Server and that server determining that the site name is valid at that location. One skilled in the art will appreciate that other regular expressions may be used. For example regular expressions may be used which recognizes a string of alpha-numeric characters in the first site name.$